

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**



## PATENT ABSTRACTS OF JAPAN

(11) Publication number: **10133696 A**(43) Date of publication of application: **22.05.98**

(51) Int. Cl. **G10L 9/14**  
**G10L 9/18**  
**H03M 7/30**

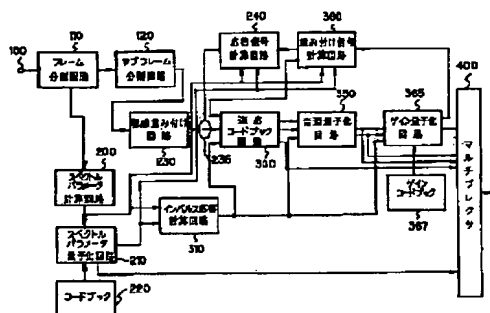
(21) Application number: **08307143**(71) Applicant: **NEC CORP**(22) Date of filing: **31.10.96**(72) Inventor: **OZAWA KAZUNORI**(54) **SPEECH ENCODING DEVICE**

## (57) Abstract:

**PROBLEM TO BE SOLVED:** To obtain excellent speech quality even when the bit rate is lowered by finding the position where predetermined conditions are met by a sound source quantization part, and searching for the best position in the search range of positions of pulses representing a sound source signal on the basis of the found position.

**SOLUTION:** A spectrum parameter calculating circuit 200 finds spectrum parameters from an input speech signal and quantizes them. An adaptive code book circuit 300 finds delay corresponding to a pitch cycle from the speech signal and calculates a pitch prediction signal to predict a pitch. Then a sound source quantizing circuit 350 constitutes a sound source signal of the speech signal with M pulses whose amplitudes are not zero, finds the sample position corresponding to a pulse position where the predetermined conditions are met for the pitch prediction signal, and sets a range of a search for the position of a pulse on the basis of a position which is shifted from the found sample position by a predetermined number of samples, thereby searching for and outputting the best position for the set range.

COPYRIGHT: (C)1998,JPO



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平10-133696

(43) 公開日 平成10年(1998) 5月22日

(51) Int.Cl.<sup>6</sup>  
G 1 0 L 9/14  
9/18  
H 0 3 M 7/30

識別記号

F I  
G 1 0 L 9/14  
9/18  
H 0 3 M 7/30

G  
J  
E  
B

審査請求 有 請求項の数12 F D (全 19 頁)

(21) 出願番号 特願平8-307143  
(22) 出願日 平成8年(1996)10月31日

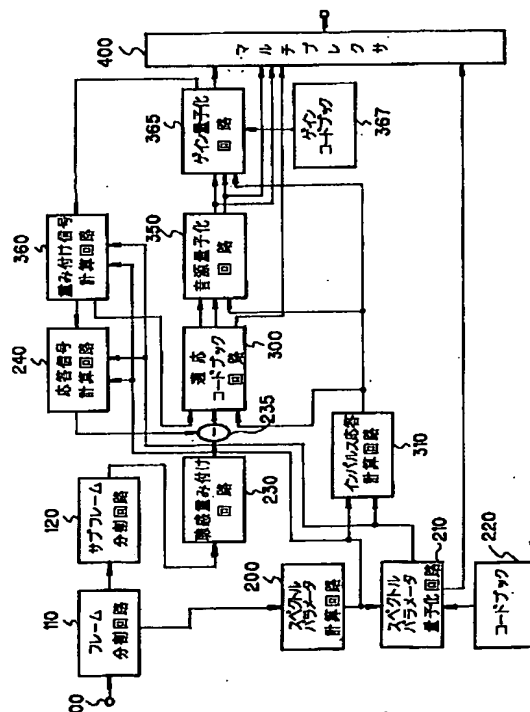
(71) 出願人 000004237  
日本電気株式会社  
東京都港区芝五丁目7番1号  
(72) 発明者 小 澤 一 範  
東京都港区芝五丁目7番1号 日本電気株  
式会社内  
(74) 代理人 弁理士 福山 正博

(54) 【発明の名称】 音声符号化装置

(57) 【要約】

【課題】 低ビットレートでも良好な音質を得る。

【解決手段】 音声符号化装置の音源量子化部360において、音源信号を個数Mの非零のパルスで構成し、スペクトルパラメータを用いて前記パルスの位置を探索する際に、Mより小さい個数毎にゲインを変化させながらパルスの位置を探索して出力する。



## 【特許請求の範囲】

【請求項1】入力音声信号からスペクトルパラメータを求めて量子化するスペクトルパラメータ計算部と、前記音声信号からピッチ周期に対応する遅延を求めてピッチ予測信号を計算しピッチ予測を行なう適応コードブック部と、前記音声信号の音源信号を個数Mの振幅が非零のパルスで構成し、前記ピッチ予測信号に対して予め定められた条件を満たす前記パルス位置対応のサンプル位置を求め、求められたサンプル位置から予め定められたサンプル数だけずらせた位置をもとに前記パルスの位置を探索する範囲を設定し、設定された範囲に対して最良の位置を探索し出力する音源量子化部とを有する音声符号化装置。

【請求項2】入力音声信号からスペクトルパラメータを求めて量子化するスペクトルパラメータ計算部と、前記音声信号からピッチ周期に当たる遅延を求めピッチ予測信号を計算しピッチ予測を行なう適応コードブック部と、前記音声信号の音源信号を個数Mの振幅が非零のパルスで構成し、先頭からピッチ周期に等しい長さの区間において前記ピッチ予測信号に対して予め定められた条件を満たすサンプル位置を求め前記位置から予め定められたサンプル数だけずらせた位置をもとにパルスの位置を探索する範囲を設定し、前記範囲に対して最良の位置を探索し出力する音源量子化部とを有する音声符号化装置。

【請求項3】入力音声信号からスペクトルパラメータを求めて量子化するスペクトルパラメータ計算部と、前記音声信号からピッチ周期に当たる遅延を求めピッチ予測信号を計算しピッチ予測を行なう適応コードブック部と、前記音声信号の音源信号を個数Mの振幅が非零のパルスで構成し、先頭からピッチ周期に等しい長さの区間において前記ピッチ予測信号に対して予め定められた条件を満たすサンプル位置を求め、前記位置から予め定められたサンプル数だけずらせた位置をもとにパルスの位置の候補を前記ピッチ周期だけずらせながら設定し、前記候補位置を探索し最良の位置を出力する音源量子化部とを有する音声符号化装置。

【請求項4】音源量子化部において、複数個のパルスの振幅もしくは極性をまとめて量子化するためのコードブックを有することを特徴とする請求項1または2または3記載の音声符号化装置。

【請求項5】入力音声信号からスペクトルパラメータを求めて量子化するスペクトルパラメータ計算部と、前記音声信号からピッチ周期に当たる遅延を求めピッチ予測信号を計算しピッチ予測を行なう適応コードブック部と、前記音声信号の音源信号を個数Mの振幅が非零のパルスで構成し、前記ピッチ予測信号に対して予め定められた条件を満たすサンプル位置を求め、複数種のずらし量の各々を用いて前記位置からずらした後の位置をもとに前記パルスの位置を探索する範囲を設定し前記範囲に

対して位置を探索し、最良となるずらし量とパルスの位置の組合せを出力する音源量子化部とを有する音声符号化装置。

【請求項6】入力音声信号からスペクトルパラメータを求めて量子化するスペクトルパラメータ計算部と、前記音声信号からピッチ周期に当たる遅延を求めピッチ予測信号を計算しピッチ予測を行なう適応コードブック部と、前記音声信号の音源信号を個数Mの振幅が非零のパルスで構成し、先頭からピッチ周期に等しい長さの区間において前記ピッチ予測信号に対して予め定められた条件を満たすサンプル位置を求め、複数種のずらし量の各々を用いて前記位置からずらせた後の位置をもとに前記パルスの位置を探索する範囲を設定し前記範囲に対して位置を探索し、最良となるずらし量とパルスの位置の組合せを出力する音源量子化部とを有する音声符号化装置。

【請求項7】入力音声信号からスペクトルパラメータを求めて量子化するスペクトルパラメータ計算部と、前記音声信号からピッチ周期に当たる遅延を求めピッチ予測信号を計算しピッチ予測を行なう適応コードブック部と、前記音声信号の音源信号を個数Mの振幅が非零のパルスで構成し、先頭からピッチ周期に等しい長さの区間において前記ピッチ予測信号に対して予め定められた条件を満たすサンプル位置を求め、複数種のずらし量の各々を用いて前記位置からずらせた後の位置をもとに、更に前記ピッチ周期だけずらせながら前記パルスをたてる位置の候補を設定し、前記位置を探索し、最良となるずらし量とパルスの位置の組合せを出力する音源量子化部とを有する音声符号化装置。

【請求項8】音源量子化部において、複数個のパルスの振幅もしくは極性をまとめて量子化するためのコードブックを有することを特徴とする請求項5または6または7記載の音声符号化装置。

【請求項9】入力音声信号からスペクトルパラメータを求めて量子化するスペクトルパラメータ計算部と、入力音声信号から特徴量を抽出して複数のモードを判別し出力するモード判別部と、前記音声信号からピッチ周期に当たる遅延を求めピッチ予測信号を計算しピッチ予測を行なう適応コードブック部と、前記音声信号の音源信号を個数Mの振幅が非零のパルスで構成し、予め定められたモードの場合に、前記ピッチ予測信号に対して予め定められた条件を満たすサンプル位置を求め、前記位置をもとに、前記パルスの位置を探索する範囲を設定し、前記範囲に対して最良を探索し出力する音源量子化部とを有することを特徴とする音声符号化装置。

【請求項10】前記特徴量は平均ピッチ予測ゲインである請求項9に記載の音声符号化装置。

【請求項11】前記モード判別部は前記平均ピッチ予測ゲインと予め定められた複数個のしきい値との比較結果に基づいてモードを判別する請求項9に記載の音声符号

化装置。

【請求項12】入力音声信号からスペクトルパラメータを求めて量子化するスペクトルパラメータ計算部と、前記音声信号からピッチ周期に対応する遅延を求めてピッチ予測信号を計算し、ピッチ予測を行なう適応コードブック部と、前記適応コードブックで求めたピッチ予測信号に対して予め定められた条件を満たす位置を求め、求められた位置に基づいて音源信号を表わす複数のパルスの位置の探索範囲を設定し、この探索範囲の中で前記複数のパルスの最良の位置を探索する音源量子化部とを備えて成ることを特徴とする音声符号化装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、低ビットレート且つ高品質で音声信号を符号化する音声符号化装置に関する。

【従来の技術】

【0002】音声信号を高効率に符号化する方式としては、例えば、M.Schroeder and B.Atal氏による"Code-excited linear prediction:High quality speech at very low bit rates"(Proc. ICASSP, pp. 937-940, 1985年)と題した論文(文献1)や、Klejin氏らによる"Improved speech quality and efficient vector quantization in CELP"(Proc. ICASSP, pp. 155-158, 1988年)と題した論文(文献2)等に記載されているCELP(Code Excited Linear Predictive Coding)が知られている。この従来例では、送信側では、フレーム毎(例えば20ms)に音声信号から線形予測(LPC)分析を用いて、音声信号のスペクトル特性を表わすスペクトルパラメータを抽出する。フレームを更にサブフレーム(例えば5ms)に分割し、サブフレーム毎に過去の音源信号を基に適応コードブックにおけるパラメータ(ピッチ周期に対応する遅延パラメータとゲインパラメータ)を抽出し、適応コードブックにより前記サブフレームの音声信号をピッチ予測する。ピッチ予測して求めた音源信号に対して、予め定められた種類の雑音信号からなる音源コードブック(ベクトル量子化コードブック)から最適な音源コードベクトルを選択し、最適なゲインを計算することにより、音源信号を量子化する。音源コードベクトルの選択の仕方は、選択した雑音信号により合成した信号と、前記残差信号との誤差電力を最小化するように行う。そして、選択されたコードベクトルの種類を表わすインデックスとゲインならびに、前記スペクトルパラメータと適応コードブックのパラメータをマルチプレクサ部により組み合わせて伝送する。受信側の動作、構成は周知であるので説明は省略する。

【0003】

【発明が解決しようとする課題】しかしながら、前記の音声符号化装置では、音源コードブックから最適な音源コードベクトルを選択するのに多大な演算量を要すると

いう問題があった。これは、文献1や2の方法では、音源コードベクトルを選択するのに、各コードベクトルに対してフィルタリングもしくは畳み込み演算を、コードブックに格納されているコードベクトルの個数だけ繰り返さなければならないことに起因する。例えば、コードブックのビット数がBビットで次元数がNのときは、フィルタリングあるいは畳み込み演算の時のフィルタあるいはインパルス応答長をKとすると、演算量は1秒当たり、 $N \times K \times 2^B \times 8000 / N$ だけ必要となる。一例として、 $B=10$ 、 $N=40$ 、 $K=10$ とすると、1秒当たり81,920,000回の演算が必要となり、極めて膨大な演算量になってしまうという問題点があった。

【0004】そこで、音源コードブック探索に必要な演算量を低減する方法として、種々のものが提案されている。例えば、ACELP(Algebraic Code Excited Linear Prediction)方式が、例えば、C.Laflammeらによる"16 kbps wideband speech coding technique based on algebraic CELP"と題した論文(Proc. ICASSP, pp. 13-16, 1991)(文献3)等に関示されている。ACELP方式によれば、音源信号を複数のパルスで表わし、各パルスのたつ位置は、各パルス毎に予め定められた位置の候補から選択し、これを予め定められたビット数で表わして伝送する。ここで、各パルスの振幅は+1.0もしくは-1.0に限定されているため、パルス探索の演算量を大幅に低減化できる。

【0005】文献3の従来方式では、演算量を大幅に低減化することが可能となるが、ビットレートを低減化すると、サブフレーム当たりのパルスの個数が急速に減少し、音質が大幅に劣化するという問題がある。

【0006】そこで、本発明の目的は、上述の問題点を解決し、ビットレートが低い場合にも比較的少ない演算量で音質の劣化の少ない音声符号化方式を提供することにある。

【0007】

【課題を解決するための手段】前述の課題を解決するため、本発明の第1の態様による音声符号化方式は、入力音声信号からスペクトルパラメータを求めて量子化するスペクトルパラメータ計算部と、前記音声信号からピッチ周期に対応する遅延を求めてピッチ予測信号を計算し、ピッチ予測を行なう適応コードブック部と、前記音声信号の音源信号を個数Mの振幅が非零のパルスから構成し、前記ピッチ予測信号に対して予め定められた条件を満たす前記パルス位置対応のサンプル位置を求め、求められたサンプル位置から予め定められたサンプル数だけずらせた位置をもとに前記パルスの位置を探索する範囲を設定し、設定された範囲に対して最良の位置を探索し出力する音源量子化部とを有する。

【0008】また、本発明の第2の態様による音声符号化装置は、入力音声信号からスペクトルパラメータを求

10

20

30

40

50

めて量子化するスペクトルパラメータ計算部と、前記音声信号からピッチ周期に当たる遅延を求めピッチ予測信号を計算しピッチ予測を行なう適応コードブック部と、前記音声信号の音源信号を個数Mの振幅が非零のパルスで構成し、先頭からピッチ周期に等しい長さの区間において前記ピッチ予測信号に対して予め定められた条件を満たすサンプル位置を求め前記位置から予め定められたサンプル数だけずらせた位置をもとにパルスの位置を探索する範囲を設定し、前記範囲に対して最良の位置を探索し出力する音源量子化部とを有する。

【0009】本発明の第3の態様による音声符号化装置は、入力音声信号からスペクトルパラメータを求めて量子化するスペクトルパラメータ計算部と、前記音声信号からピッチ周期に当たる遅延を求めピッチ予測信号を計算しピッチ予測を行なう適応コードブック部と、前記音声信号の音源信号を個数Mの振幅が非零のパルスで構成し、先頭からピッチ周期に等しい長さの区間において前記ピッチ予測信号に対して予め定められた条件を満たすサンプル位置を求め、前記位置から予め定められたサンプル数だけずらせた位置をもとにパルスの位置の候補を前記ピッチ周期だけずらせながら設定し、前記候補位置を探索し最良の位置を出力する音源量子化部とを有する。

【0010】ここで、音源量子化部において、複数のパルスの振幅もしくは極性をまとめて量子化するためのコードブックを有する。

【0011】本発明の第4の態様による音声符号化装置は、入力音声信号からスペクトルパラメータを求めて量子化するスペクトルパラメータ計算部と、前記音声信号からピッチ周期に当たる遅延を求めピッチ予測信号を計算しピッチ予測を行なう適応コードブック部と、前記音声信号の音源信号を個数Mの振幅が非零のパルスで構成し、前記ピッチ予測信号に対して予め定められた条件を満たすサンプル位置を求め、複数のずらし量の各々を用いて前記位置からずらした後の位置をもとに前記パルスの位置を探索する範囲を設定し前記範囲に対して位置を探索し、最良となるずらし量とパルスの位置の組合せを出力する音源量子化部とを有する。

【0012】本発明の第5の態様による音声符号化装置は、入力音声信号からスペクトルパラメータを求めて量子化するスペクトルパラメータ計算部と、前記音声信号からピッチ周期に当たる遅延を求めピッチ予測信号を計算しピッチ予測を行なう適応コードブック部と、前記音声信号の音源信号を個数Mの振幅が非零のパルスで構成し、先頭からピッチ周期に等しい長さの区間において前記ピッチ予測信号に対して予め定められた条件を満たすサンプル位置を求め、複数のずらし量の各々を用いて前記位置からずらせた後の位置をもとに前記パルスの位置を探索する範囲を設定し前記範囲に対して位置を探索し、最良となるずらし量とパルスの位置の組合せを出力

する音源量子化部とを有する。

【0013】本発明の第6の態様による音声符号化装置は、入力音声信号からスペクトルパラメータを求めて量子化するスペクトルパラメータ計算部と、前記音声信号からピッチ周期に当たる遅延を求めピッチ予測信号を計算しピッチ予測を行なう適応コードブック部と、前記音声信号の音源信号を個数Mの振幅が非零のパルスで構成し、先頭からピッチ周期に等しい長さの区間において前記ピッチ予測信号に対して予め定められた条件を満たすサンプル位置を求め、複数のずらし量の各々を用いて前記位置からずらせた後の位置をもとに、更に前記ピッチ周期だけずらせながら前記パルスをたてる位置の候補を設定し、前記位置を探索し、最良となるずらし量とパルスの位置の組合せを出力する音源量子化部とを有する。

【0014】ここで、音源量子化部において、複数のパルスの振幅もしくは極性をまとめて量子化するためのコードブックを有する。

【0015】本発明の第7の態様による音声符号化装置は、入力音声信号からスペクトルパラメータを求めて量子化するスペクトルパラメータ計算部と、入力音声信号から特徴量を抽出して複数のモードを判別し出力するモード判別部と、前記音声信号からピッチ周期に当たる遅延を求めピッチ予測信号を計算しピッチ予測を行なう適応コードブック部と、前記音声信号の音源信号を個数Mの振幅が非零のパルスで構成し、予め定められたモードの場合に、前記ピッチ予測信号に対して予め定められた条件を満たすサンプル位置を求め、前記位置をもとに、前記パルスの位置を探索する範囲を設定し、前記範囲に対して最良を探索し出力する音源量子化部とを有する。

【0016】ここで、前記特徴量は平均ピッチ予測ゲインであり、また前記モード判別部は前記平均ピッチ予測ゲインと予め定められた複数のしきい値との比較結果に基づいてモードを判別する。

【0017】本発明の第8の態様による音声符号化装置は、入力音声信号からスペクトルパラメータを求めて量子化するスペクトルパラメータ計算部と、前記音声信号からピッチ周期に対応する遅延を求めてピッチ予測信号を計算し、ピッチ予測を行なう適応コードブック部と、前記適応コードブックで求めたピッチ予測信号に対して予め定められた条件を満たす位置を求め、求められた位置に基づいて音源信号を表わす複数のパルスの位置の探索範囲を設定し、この探索範囲の中で前記複数のパルスの最良の位置を探索する音源量子化部とを備えて成る。

【0018】

【実施態様】図1は本発明による音声符号化装置の第1の実施の形態を示すブロック図である。図1において、入力端子100から音声信号が入力され、フレーム分割回路110では上記音声信号がフレーム（例えば10m

s) 毎に分割され、サブフレーム分割回路120では、上記フレーム音声信号をフレームよりも短いサブフレーム(例えば、5ms)に分割される。

【0019】スペクトルパラメータ計算回路200は、少なくとも一つのサブフレームの音声信号に対して、サブフレーム長よりも長い窓(例えば、24ms)をかけて音声信号を切り出してスペクトルパラメータを予め定められた次数(例えばP=10次)計算する。ここで、スペクトルパラメータの計算には、周知のLPC分析や、Burg分析等を用いることができる。ここでは、Burg分析を用いることとする。Burg分析の詳細については、中溝著による“信号解析とシステム同定”と題した単行本(コロナ社1988年刊)の82~87頁(文献4)等に記載されているので説明は省略する。更に、スペクトルパラメータ計算部は、Burg法により計算された線形予測係数 $\alpha_i$  ( $i=1, \dots, 10$ )を量子化や補間に適したLSPパラメータに変換する。ここで、線形予測係数からLSPへの変換は、菅村他によ

$$D_j = \sum_i^P W(i) [LSP(i) - QLSP(i)]^2 \quad (1)$$

ここで、LSP(i)、QLSP(i)、W(i)はそれぞれ、量子化前のi次目のLSP、コードブック220に格納されたj番目のコードベクトル、重み係数である。

【0021】以下では、量子化法として、ベクトル量子化を用いるものとし、第2サブフレームのLSPパラメータを量子化するものとする。LSPパラメータのベクトル量子化の手法としては周知の手法を用いることができる。具体的な手法は、例えば、特開平4-171500号公報(特願平5-297600号)(文献6)、特開平4-363000号公報(特願平3-261925号)(文献7)、特開平5-6199号公報(特願平3-155049号)(文献8)、T.Nomura et al., による“LSP Coding Using VQSVQ with Interpolation in 4.075kbps M-LCELP Speech Coder”と題した論文(Proc. Mobile Multimedia Communications, pp. B. 2. 5, 1993)(文献9)等を参照できるのでここでは説明は略する。

【0022】また、スペクトルパラメータ量子化回路210は、第2サブフレームで量子化したLSPパラメータをもとに、第1サブフレームのLSPパラメータを復元する。ここでは、現フレームの第2サブフレームの量子化LSPパラメータと1つ過去のフレームの第2サブフレームの量子化LSPを直線補間して、第1サブフレームのLSPを復元する。ここで、量子化前のLSPと量子化後のLSPとの誤差電力を量子化するコードベク

\*る”線スペクトル対(LSP)音声分析合成方式による音声情報圧縮”と題した論文(電子通信学会論文誌、J64-A, pp. 599-606, 1981年)(文献5)を参照することができる。例えば、第2サブフレームでBurg法により求めた線形予測係数を、LSPパラメータに変換し、第1サブフレームのLSPを直線補間により求めて、第1サブフレームのLSPを逆変換して線形予測係数に戻し、第1、2サブフレームの線形予測係数 $\alpha_n$ ,  $i=1, \dots, 10$ ,  $l=1, \dots, 2$ )を聴感重み付け回路230に出力する。また、第2サブフレームのLSPをスペクトルパラメータ量子化回路210へ出力する。

【0020】スペクトルパラメータ量子化回路210は、予め定められたサブフレームのLSPパラメータをコードブック220を用いて効率的に量子化し、下式の歪みを最小化する量子化値を出力する。

【数1】

※トルを1種類選択した後に、直線補間により第1サブフレームのLSPを復元できる。

【0023】以上により復元した第1サブフレームのLSPと第2サブフレームの量子化LSPをサブフレーム毎に線形予測係数 $\alpha_n'$  ( $i=1, \dots, 10$ ,  $l=1, \dots, 2$ )に変換し、インパルス応答計算回路310へ出力する。また、第2サブフレームの量子化LSPのコードベクトルを表わすインデックスをマルチプレクサ400に出力する。

【0024】聴感重み付け回路230は、スペクトルパラメータ計算回路200から、各サブフレーム毎に量子化前の線形予測係数 $\alpha_{ij}'$  ( $i=1, \dots, P$ )を入力し、前記文献1に基づき、サブフレームの音声信号に対して聴感重み付けを行い、聴感重み付け信号を出力する。

【0025】応答信号計算回路240は、スペクトルパラメータ計算回路200から、各サブフレーム毎に線形予測係数 $\alpha_i$ を入力し、スペクトルパラメータ量子化回路210から、量子化、補間して復元した線形予測係数 $\alpha_i'$ をサブフレーム毎に入力し、保存されているフィルタメモリの値を用いて、入力信号を $d(n)=0$ とした応答信号を1サブフレーム分計算し、減算器235へ出力する。ここで、応答信号 $x_z(n)$ は下式で表される。

【数2】

$$x_z(n) = d(n) - \sum_{i=1}^P \alpha_i d(n-i) + \sum_{i=1}^P \alpha_i \gamma^i y(n-i) + \sum_{i=1}^P \alpha_i' \gamma^i x_z(n-i) \quad (2)$$

但し、 $n-i \leq 0$  のときは

\* \* 【数3】

$$y(n-i) = p(N + (n-i)) \quad (3)$$

【数4】

$$x_z(n-i) = s_w(N + (n-i)) \quad (4)$$

ここで、 $N$ はサブフレーム長を示す。 $\gamma$ は、聴感重み付け量を制御する重み係数であり、下記の式(6)と同一の値である。 $s_w(n)$ 、 $p(n)$ は、それぞれ、重み付け信号計算回路の出力信号、後述の式(6)における

※示す。

【0026】減算器235は、下式により、聴感重み付け信号から応答信号を1サブフレーム分減算し、 $x_w'(n)$ を適応コードブック回路300へ出力する。

【数5】

$$x_w'(n) = x_w(n) - x_z(n) \quad (5)$$

【0027】インパルス応答計算回路310は、 $z$ 変換が下式で表される聴感重み付けフィルタのインパルス応答  $h_w(n)$  を予め定められた点数  $L$  だけ計算し、適応

★コードブック回路300、音源量子化回路350へ出力する。

【数6】

$$H_w(z) = \frac{1 - \sum_{i=1}^P \alpha_i z^{-i}}{1 - \sum_{i=1}^P \alpha_i \gamma^i z^{-i}} \frac{1}{1 - \sum_{i=1}^P \alpha_i' \gamma^i z^{-i}} \quad (6)$$

【0028】適応コードブック回路300は、重み付け信号計算回路360から退去の音源信号  $v(n)$  を、減算器235から出力信号  $x_w'(n)$  を、インパルス応答計算回路310から聴感重み付けインパルス応答  $h_w$

☆  $v(n)$  を入力する。ピッチ周期に対応する遅延  $T$  を下式の歪みを最小化するように求め、遅延を表わすインデックスをマルチプレクサ400に出力する。

【数7】

$$D_T = \sum_{n=0}^{N-1} x_w'^2(n) - \left[ \sum_{n=0}^{N-1} x_w'(n) y_w(n-T) \right]^2 / \left[ \sum_{n=0}^{N-1} y_w^2(n-T) \right] \quad (7)$$

ここで、

◆ ◆ 【数8】

$$y_w(n-T) = v(n-T) * h_w(n) \quad (8)$$

はピッチ予測信号を示し、記号\*は畳み込み演算を表わす。ゲイン  $\beta$  は下式に従い求める。

【数9】

$$\beta = \sum_{n=0}^{N-1} x_w'(n) y_w(n-T) / \sum_{n=0}^{N-1} y_w^2(n-T) \quad (9)$$

【0029】ここで、女性音や、子供の声に対して、遅延の抽出精度を向上させるために、遅延を整数サンプルではなく、小数サンプル値で求めてもよい。具体的な方法は、例えば、P. Kroonらによる、“Pitch predictors with high temporal resolution”と題した論文(Proc. I CASSP, pp. 661-664, 1990年) (文献10)等を参照する

ことができる。

【0030】更に、適応コードブック回路300は、選択された遅延とゲインを用いて下式に従いピッチ予測を行ない、予測残差信号  $z_w(n)$  を音源量子化回路350へ出力する。

【数10】



$$z_w(n) = x'_w(n) - \beta v(n-T) * h_w(n)$$

(10)

更に、選択された遅延を用いたピッチ予測信号を音源量子化回路350へ出力する。

【0031】音源量子化回路350では、サブフレームに対して、振幅が非零のM個のパルスをたてる。

【0032】音源量子化回路350の構成を示すブロック図を図2に示す。絶対値最大位置検出回路351は、ピッチ予測信号 $y_p(n)$ に対して、予め定められた条件を満たすサンプル位置を検出する。ここでは、「振幅の絶対値が最大」という条件を使用し、それを満たすサンプル位置を検出し、位置探索範囲設定回路352へ出力する。

【0033】位置探索範囲設定回路352は、入力したサンプル位置に対して予め定められた固定のサンプル数だけ未来あるいは過去にずらした後に、各パルスの位置の探索範囲を設定する。

\*【0034】例えば、入力したサンプル位置をDとし、5msサブフレーム(40サンプル)に5個のパルスを求める例を考えると、各パルスの探索範囲に含まれる位置の候補の例は下表のようになる。

第1パルス D-L, D-L+5, . . .

第2パルス D-L+1, D-L+6, . . .

10 第3パルス D-L+2, D-L+7, . . .

第4パルス D-L+3, D-L+8, . . .

第5パルス D-L+4, D-L+9, . . .

【0035】次に、 $z_p(n)$ 、 $h_p(n)$ を入力し、第1の相関関数計算回路353、第2の相関関数計算回路354は、それぞれ、下式に従い、第1の相関関数 $d(n)$ 、第2の相関関数 $\phi$ を計算する。

【数11】

$$d(n) = \sum_{i=n}^{N-1} z_w(i) h_w(i-n), \quad n=0, \dots, N-1 \quad (11)$$

【数12】

$$\phi(p, q) = \sum_{n=\max(p, q)}^{N-1} h_w(n-p) h_w(n-q), \quad p, q=0, \dots, N-1 \quad (12)$$

【0036】パルス極性設定回路355は、位置探索範囲設定回路352で設定された探索範囲における各パルスの候補位置に対して、第1の相関関数 $d(n)$ の極性を抽出し出力する。

【0037】パルス位置探索回路356は、上表に示した候補位置の組合せに対して次式を計算し、次式を最大化する位置を最適位置として選択する。

$$C_k = \sum_{k=1}^M \text{sign}(k) d(m_k) \quad (14)$$

【数15】

$$E = \sum_{k=1}^M \text{sign}(k)^2 \phi(m_k, m_k) + 2 \sum_{k=1}^{M-1} \sum_{i=k+1}^M \text{sign}(k) \text{sign}(i) \phi(m_k, m_i) \quad (15)$$

である。ここで、 $\text{sign}(k)$ は、k番目のパルスの極性を示し、パルス極性設定回路355にて予め抽出したものを使用する。以上により、M個のパルスの極性と位置がゲイン量子化回路365に出力される。

【0038】また、パルスの位置を予め定められたビット数で量子化し、位置を表わすインデックスをマルチプレクサに出力する。また、パルスの極性をマルチプレクサ

【数13】

$$D = \frac{C_k^2}{E_k} \quad (13)$$

ここで、パルスの個数をMとすると、

【数14】

※400に出力する。

【0039】ゲイン量子化回路365は、ゲインコードブック367からゲインコードベクトルを読み出し、選択された位置に対して、下式を最小化するゲインコードベクトルを選択し、最終的に歪みを最小化する振幅コードベクトルとゲインコードベクトルの組合せを選択する。

※50

【0040】ここでは、適応コードブックのゲイン $\beta'$ と、パルスで表わした音源のゲイン $G'$ の2種のゲイン\*

$$D_t = \sum_{n=0}^{N-1} [x_w(n) - \beta'_t v(n-T) * h_w(n) - G'_t \sum_{k=1}^M \text{sign}(k) h_w(n-m_k)]^2 \quad (16)$$

ここで、 $\beta'_t$ 、 $G'_t$ は、ゲインコードブック367に格納された2次元ゲインコードベクトルにおけるt番目の要素である。上式の計算を、ゲインコードベクトルの各々に対して繰り返し、歪み $D_t$ を最小化するゲインコードベクトルを選択する。選択されたゲインコードベクトルを表わすインデックスをマルチプレクサ400に出力※

$$v(n) = \beta'_t v(n-T) + G'_t \sum_{k=1}^M \text{sign}(k) \delta(n-m_k) \quad (17)$$

$v(n)$ は適応コードブック回路300に出力される。

【0042】次に、スペクトルパラメータ計算回路200の出力パラメータ、スペクトルパラメータ量子化回路210の出力パラメータを用いて下式により、応答信号★

$$s_w(n) = v(n) - \sum_{i=1}^P a_i v(n-i) + \sum_{i=1}^P a_i \gamma^i p(n-i) + \sum_{i=1}^P a'_i \gamma^i s_w(n-i) \quad (18)$$

【0043】第2の実施の形態を示すブロック図を図3に示す。ここでは、音源量子化回路450の動作が図1と異なる。

【0044】音源量子化回路450の構成を図4に示す。音源量子化回路450は、予測信号 $y_p(n)$ 、予測残差信号 $z_p(n)$ 、聴感重み付けインパルス応答 $h_p(n)$ のみならず、適応コードブックの遅延 $T$ を入力する。

【0045】絶対値最大位置計算回路451は、ピッチ周期に相当する遅延 $T$ を入力し、ピッチ予測信号 $y_p(n)$ に対して、サブフレームの先頭から $T$ サンプルまでの範囲で絶対値を最大にするサンプル位置を検出し、位置探索範囲設定回路352に出力する。

【0046】第3の実施の形態を示すブロック図を図5に示す。ここでは、音源量子化回路500の動作が図3と異なる。音源量子化回路550の構成図を図6に示す。

【0047】位置探索範囲設定回路552は、入力したサンプル位置に対して予め定められた固定のサンプル数 $L$ だけ未来あるいは過去にずらした位置を基点とし、遅延 $T$ だけずらしながら、各パルスの位置の候補を設定し、パルス位置探索回路356に出力する。

【0048】例えば、入力したサンプル位置を $D$ とし、☆

$$C_j^2/E_j$$

ここで、

\*を同時にベクトル量子化する例について示す。

【数16】

※する。

【0041】重み付け信号計算回路360は、それぞれのインデックスを入力し、インデックスからそれに対応するコードベクトルを読み出し、まず下式に基づき駆動音源信号 $v(n)$ を求める。

【数17】

★ $s_p(n)$ をサブフレーム毎に計算し、応答信号計算回路240へ出力する。

【数18】

☆5msサブフレーム(40サンプル)に5個のパルスを求める例を考えると、各パルスの位置の候補の例は下表のようになる。

第1パルス  $D-L, D-L+T, \dots$   
第2パルス  $D-L+1, D-L+T, \dots$   
第3パルス  $D-L+2, D-L+T, \dots$   
第4パルス  $D-L+3, D-L+T, \dots$   
第5パルス  $D-L+4, D-L+T, \dots$

【0049】第4の実施の形態を示すブロック図を図7に示す。ここでは、第1の実施の形態において、振幅コードブックを使用する例について説明するが、第2、第3の実施の形態に対して振幅コードブックを使用する場合も同様の変更により実現できる。

【0050】図7は、図1に比べ、音源量子化回路390と振幅コードブック395が異なっている。音源量子化回路390の構成を図8に示す。振幅コードブック395を用いてパルスの振幅を量子化する。

【0051】パルス位置探索回路356において $M$ 個のパルスに対して位置が求まった後で、振幅量子化回路397において、次式を最大化するように、振幅コードベクトルを振幅コードブック395から選択し、インデックスを出力する。

【数19】

$$(19)$$

◆ ◆ 【数20】

$$C_j = \sum_{k=1}^M g'_{kj} d(m_k) \quad (20)$$

【数21】

$$E_j = \sum_{k=1}^M g'^2_{kj} \phi(m_k, m_k) + 2 \sum_{k=1}^{M-1} \sum_{i=k+1}^M g'_{kj} g'_{ij} \phi(m_k, m_i) \quad (21)$$

である。ここで、 $g_{kj}$ は、 $k$ 番目のパルスの $j$ 番目の振幅コードベクトルである。

【0052】音源量子化回路390は、選択された振幅コードベクトルを表すインデックスをマルチプレクサ400に出力する。また、位置の値、振幅コードベクトルの値をゲイン量子化回路400に出力する。

【0053】なお、本実施例では、振幅コードブックを使用した、代わりに、各パルスの極性を示す極性コードブックを使用して探索してもよい。

【0054】図9は、第5の実施の形態を示すブロック図である。図において、音源量子化回路600の動作が図1と異なるので、図10を用いて構成を説明する。

【0055】図10は音源量子化回路600の構成を示すブロック図である。位置探索範囲設定回路652は、絶対値最大位置検出回路351の出力位置に対して、複数種（例えばQ種）のずらし量の各々の分だけずらした位置を基点として各パルスの探索範囲ならびに位置のセットを設定し、パルスの位置の設置のセットをずらし量の種類分だけパルス極性設定回路655とパルス位置探索回路656に出力する。

【0056】パルス極性設定回路655は、位置探索回路652の複数種の候補位置の各々に対して極性を抽出し、パルス位置探索回路656へ出力する。

【0057】パルス位置探索回路656は、複数種の候補位置の各々に対して、第1の相関関数、第2の相関関数、極性を用いて、式(13)を最大化する位置を探索する。この処理をずらしの種類であるQ回繰り返し、Q種の中で、式(13)を最大化する位置を最終的に選択し、各パルスの位置と、ずらし量とを出力する。なお、ずらし量はマルチプレクサ400に出力される。

【0058】図11は、第6の実施の形態を示すブロック図である。図において、音源量子化回路650の動作が図3と異なるので、図12を用いて構成を説明する。

【0059】図12は音源量子化回路650の構成を示すブロック図である。位置探索範囲設定回路652は、絶対値最大位置検出回路451の出力位置に対して、複数種（例えばQ種）のずらし量の各々の分だけずらした位置を基点として、各パルスの位置を設定し、パルスの位置のセットをずらし量の種類分だけパルス極性設定回路655とパルス位置探索回路656に出力する。

\*【0060】パルス極性設定回路655は、位置探索回路652の複数種の候補位置の各々に対して極性を抽出し、パルス位置探索回路656へ出力する。

【0061】パルス位置探索回路656は、複数種の候補位置の各々に対して、第1の相関関数、第2の相関関数、極性を用いて、式(13)を最大化する位置を探索する。この処理をずらしの種類であるQ回繰り返し、Q種の中で、式(13)を最大化する位置を最終的に選択し、各パルスの位置と、ずらし量とを出力する。なお、ずらし量はマルチプレクサ400に出力される。

【0062】図13は、第7の実施の形態を示すブロック図である。図において、音源量子化回路750の動作が図5と異なるので、図14を用いて構成を説明する。

【0063】図14は音源量子化回路750の構成を示すブロック図である。位置探索範囲設定回路752は、絶対値最大位置検出回路451の出力位置に対して、複数種（例えば、Q種）のずらし量の各々の分だけずらした位置を基点として、更に遅延Tだけずらしながら各パルスの位置を設定する。このようにして各パルスの位置のセットをQ種類分パルス極性設定回路655とパルス位置探索回路656に出力する。

【0064】パルス極性設定回路655は、位置探索回路652の複数種の候補位置の各々に対して極性を抽出し、パルス位置探索回路656へ抽出する。

【0065】パルス位置探索回路656は、複数種の候補位置の各々に対して、第1の相関関数、第2の相関関数、極性を用いて、式(13)を最大化する位置を探索する。この処理をずらしの種類であるQ回繰り返し、Q種の中で、式(13)を最大化する位置を最終的に選択し、各パルスの位置と、ずらし量とを出力する。なお、ずらし量はマルチプレクサ400に出力される。

【0066】図15は、第8の実施の形態を示すブロック図である。ここでは、第5の実施例の形態を示すブロック図に、パルスの振幅を量子化する振幅コードブックを付加する例について示すが、第6、第7の実施の形態に付加することもできる。

【0067】図において、音源量子化回路850の動作が図7と異なるので、音源量子化回路850の構成を図16を用いて説明する。

\*【0068】図16は音源量子化回路85の構成を示す

ブロック図である。位置探索範囲設定回路652は、絶対値最大位置検出回路351の出力位置に対して、複数種（例えばQ種）のずらし量の各々の分だずらした位置を基点として、各パルスの位置を設定し、パルスの位置のセットをずらし量の種類分だけパルス極性設定回路655とパルス位置探索回路656に出力する。

【0069】パルス極性設定回路655は、位置探索回路652の複数種の候補位置の各々に対して極性を抽出し、パルス位置探索回路656へ出力する。

【0070】パルス位置探索回路656は、複数種の候補位置の各々に対して、第1の相関関数、第2の相関関数、極性を用いて、式(13)を最大化する位置を探索する。この処理をずらしの種類であるQ回繰り返し、Q種の中で、式(13)を最大化する位置を最終的に選択し、各パルスの位置と、ずらし量とを出力する。なお、\*

$$G = 10 \log_{10} [1/L \sum_{i=1}^L (P_i/E_i)] \quad (22)$$

ここで、Lはフレームに含まれるサブフレームの個数である。P<sub>i</sub>、E<sub>i</sub>はそれぞれ、i番目のサブフレームでの

※音声パワー、ピッチ予測誤差パワーを示す。

【数23】

$$P_i = \sum_{n=0}^{N-1} x_{wi}^2(n) \quad (23)$$

【数24】

$$E_i = P_i - [\sum_{n=0}^{N-1} x_{wi}(n)x_{wi}(n-T)]^2 / [\sum_{n=0}^{N-1} x_{wi}^2(n-T)] \quad (24)$$

ここで、Tは予測ゲインを最大化する最適遅延である。

【0073】フレーム平均ピッチ予測ゲインGを予め定められた複数個のしきい値と比較して複数種類（例えばR種）のモードに分類する。モードの個数Rとしては、例えば4を用いることが出来る。

【0074】適応コードブック回路950は、モード情報を受け取り、予め定められたモードの場合に、図1の適応コードブック回路300と同一の動作を行い、遅延、適応コードブック予測信号、予測残差信号を出力する。その他のモードに対しては、減算器235からの入力信号をそのまま出力する。

【0075】音源量子化回路960は、モード情報を受け取り、予め定められたモードの際に図1の音源量子化回路350と同一の動作を行う。

【0076】ゲイン量子化回路965は、モード情報を入力し、モード毎に設計された複数種のゲインコードブック367<sub>1</sub>から367<sub>r</sub>を切り替えてゲイン量子化に使用する。ゲイン量子化の動作は図1のゲイン量子化回路365と同一である。

【0077】上述した実施形態例に限らず、種々の変形 ★50

\*ずらし量はマルチプレクサ400に出力される。振幅量子化回路397は図8と同一の動作を行なう。

【0071】図17は、第9の実施の形態を示すブロック図である。ここでは、第1の実施の形態をもとにする例について示すが、他の実施の形態をもとにすることもできる。

【0072】モード判別回路900は、聴感重み付け回路230からフレーム単位で聴感重み付け信号を受け取り、モード判別情報を適応コードブック回路950、音源量子化回路960、ゲイン量子化回路965とマルチプレクサ400へ出力する。ここでは、モード判別に、現在のフレームの特徴量を用いる。特徴量としては、例えば、フレームで平均したピッチ予測ゲインを用いる。ピッチ予測ゲインの計算は例えば下式を用いる。

【数22】

★が可能である。例えば、複数パルスの振幅を量子化するためのコードブックを、音声信号を用いて予め学習して格納しておくこともできる。コードブックの学習法は、例えば、Linde氏らによる“An algorithm for vector quantization design”と題した論文（IEEE Trans. Commun., pp. 84-95, January, 1980）（文献11）等を参照できる。

【0078】振幅コードブックの代わりに、パルスの個数に等しいビット数だけ各パルスの極性の組み合わせを用意した極性コードブックを有するようにしてもよい。

【0079】

【発明の効果】以上説明したように、本発明によれば、音源量子化部において、適応コードブックで求めたピッチ予測信号に対して予め定められた条件を満たす位置を求め、前記位置を基に、音源信号を表わす複数個のパルスの位置の探索範囲を設定し、この範囲の中で最良の位置を探索する。これにより、パルスの位置の探索範囲をピッチ波形に同期させて、ピッチ波形を表わすための音源信号を良好に表わすことが出来るので、ビットレートを低減化しても、従来方式に比べ良好な音質が得られ

る。

【0081】更に本発明によれば、入力音声から特徴量を抽出して複数のモードを判別し、予め定められたモードにおいて、音源量子化部で上述の処理を行うことにより、音声の周期性が強いモード部分に対する音質を改善することが出来る。

#### 【図面の簡単な説明】

【図1】本発明による音声符号化装置の第1の実施形態を示す構成ブロック図である。

【図2】第1の実施の形態における音源量子化回路350の構成を示す図である。

【図3】本発明による音声符号化装置の第2の実施形態を示す構成ブロック図である。

【図4】第2の実施の形態における音源量子化回路450の構成を示す図である。

【図5】本発明による音声符号化装置の第3の実施形態を示す構成ブロック図である。

【図6】第3の実施の形態における音源量子化回路550の構成を示す図である。

【図7】本発明による音声符号化装置の第4の実施形態を示す構成ブロック図である。

【図8】第4の実施の形態における音源量子化回路390の構成を示す図である。

【図9】本発明による音声符号化装置の第5の実施形態を示す構成ブロック図である。

【図10】第5の実施の形態における音源量子化回路600の構成を示す図である。

【図11】本発明による音声符号化装置の第6の実施形態を示す構成ブロック図である。

【図12】第6の実施の形態における音源量子化回路6 \* 30

\* 50の構成を示す図である。

【図13】本発明による音声符号化装置の第7の実施形態を示す構成ブロック図である。

【図14】第7の実施の形態における音源量子化回路750の構成を示す図である。

【図15】本発明による音声符号化装置の第8の実施形態を示す構成ブロック図である。

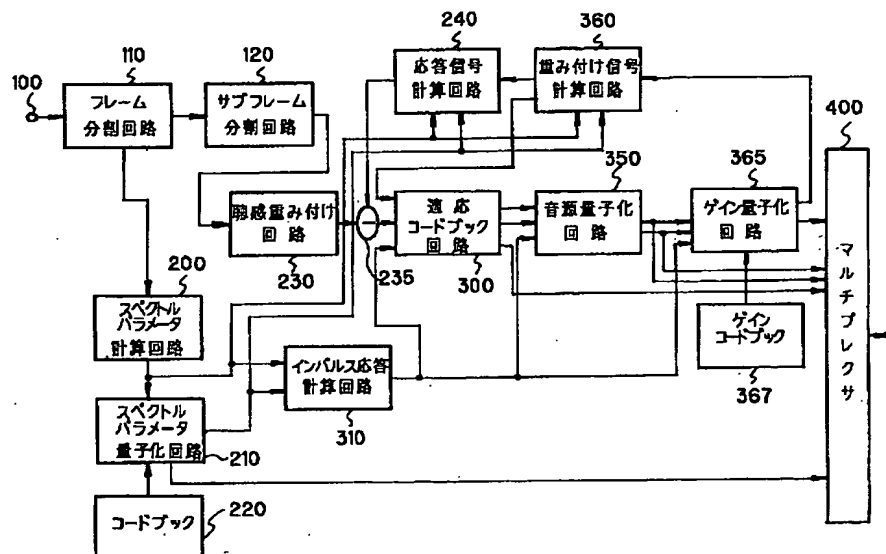
【図16】第8の実施の形態における音源量子化回路850の構成を示す図である。

【図17】本発明による音声符号化装置の第9の実施形態を示す構成ブロック図である。

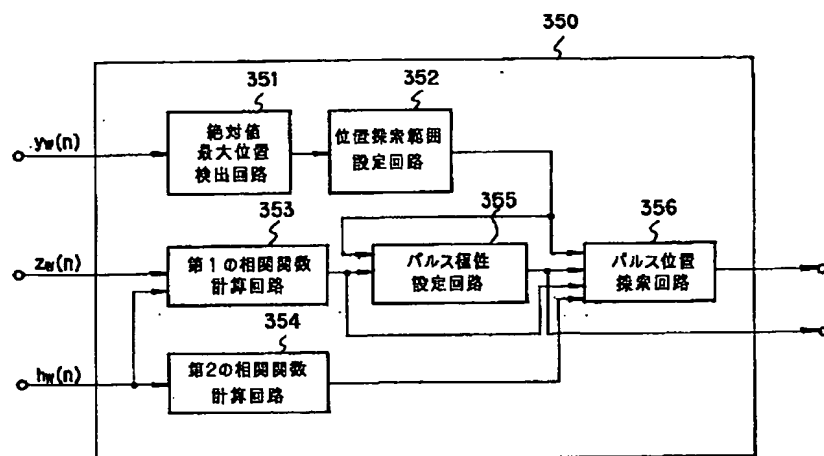
#### 【符号の説明】

- 110 フレーム分割回路
- 120 サブフレーム分割回路
- 200 スペクトルパラメータ計算回路
- 210 スペクトルパラメータ量子化回路
- 220 コードブック
- 230 聴感重み付け回路
- 235 減算回路
- 240 応答信号計算回路
- 310 インパルス応答計算回路
- 350、390、450、500、600、650、750、850、960 音源量子化回路
- 360 重み付け信号計算回路
- 365、965 ゲイン量子化回路
- 395 振幅コードブック
- 367 ゲインコードブック
- 400 マルチプレクサ
- 900 モード判別回路

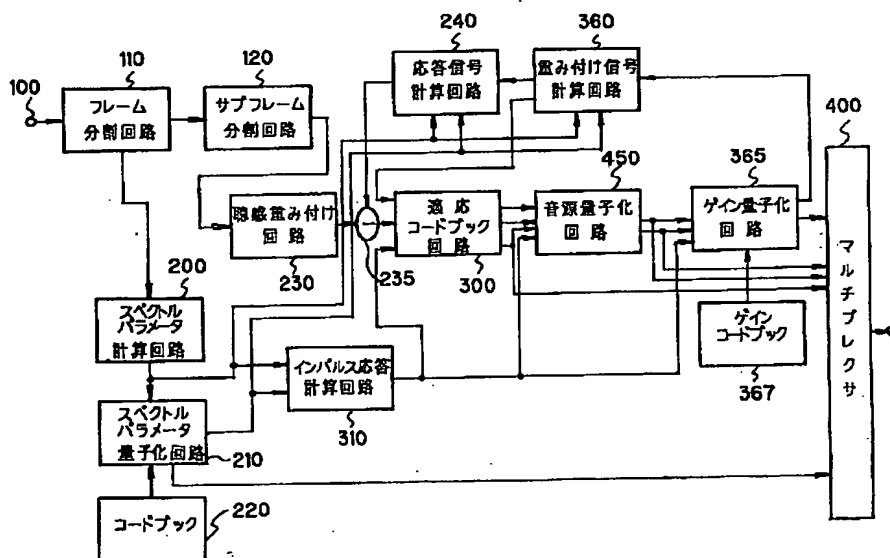
【図1】



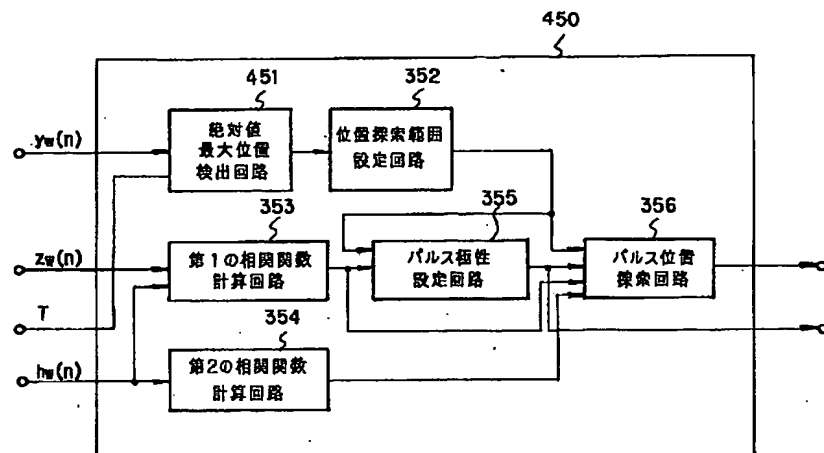
【図2】



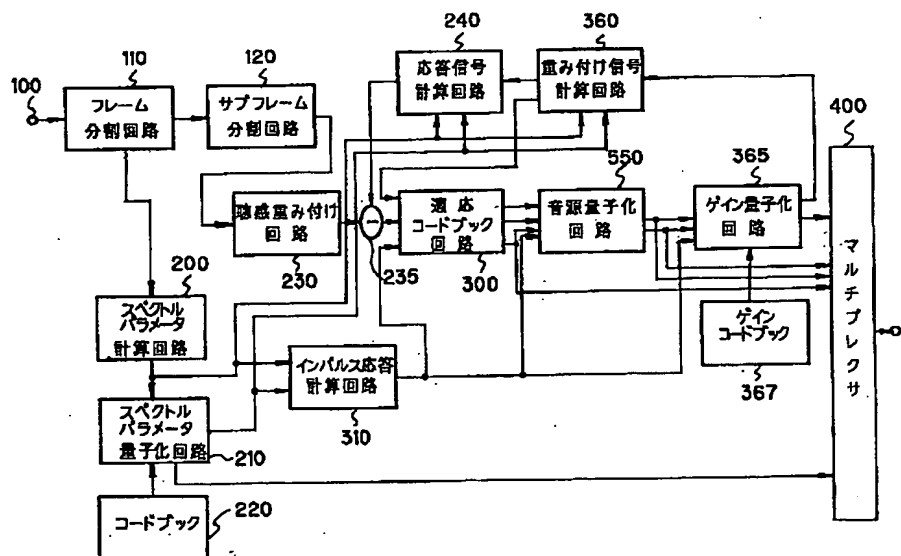
【図3】



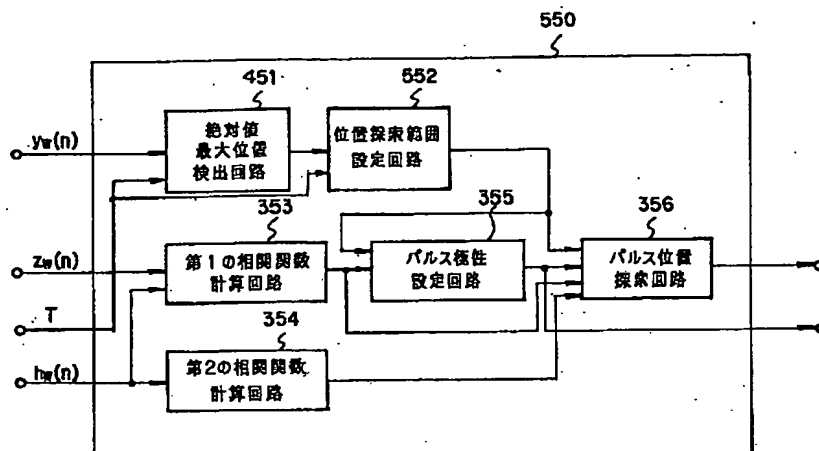
【図 4】



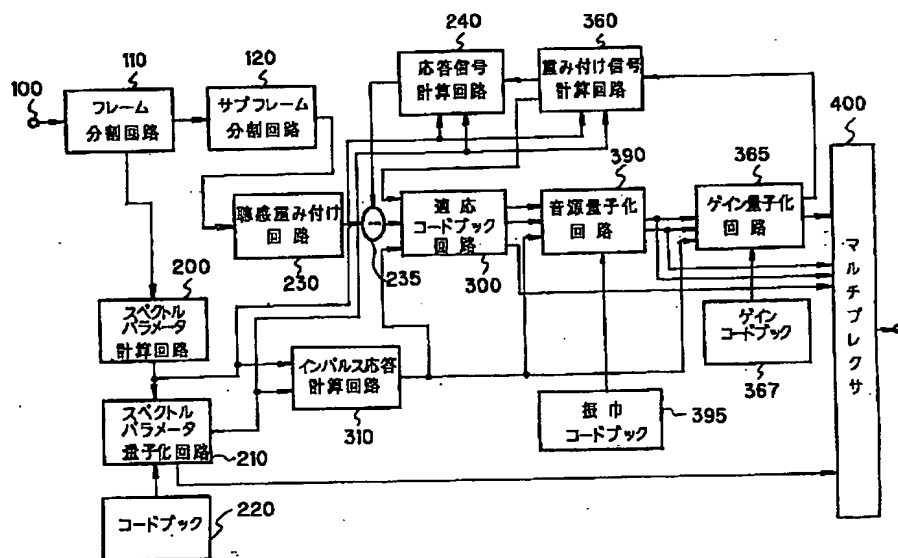
【図 5】



【図6】

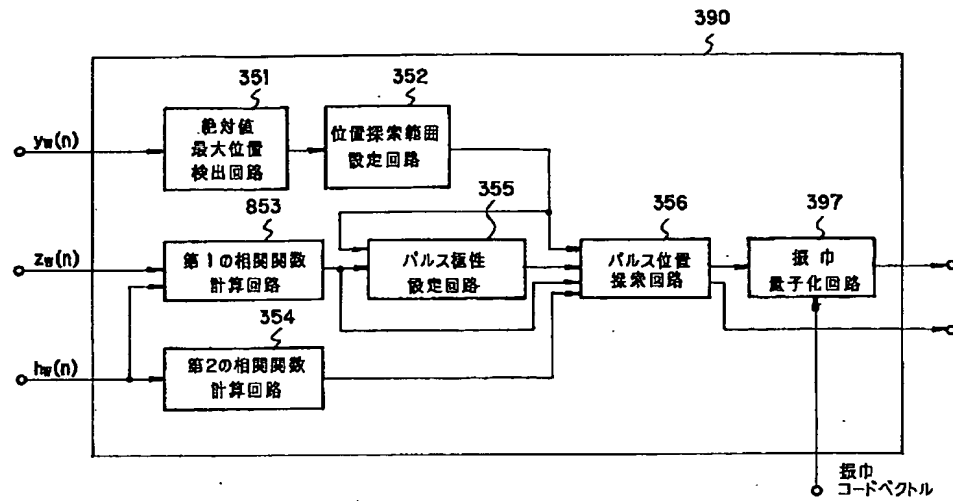


【図7】

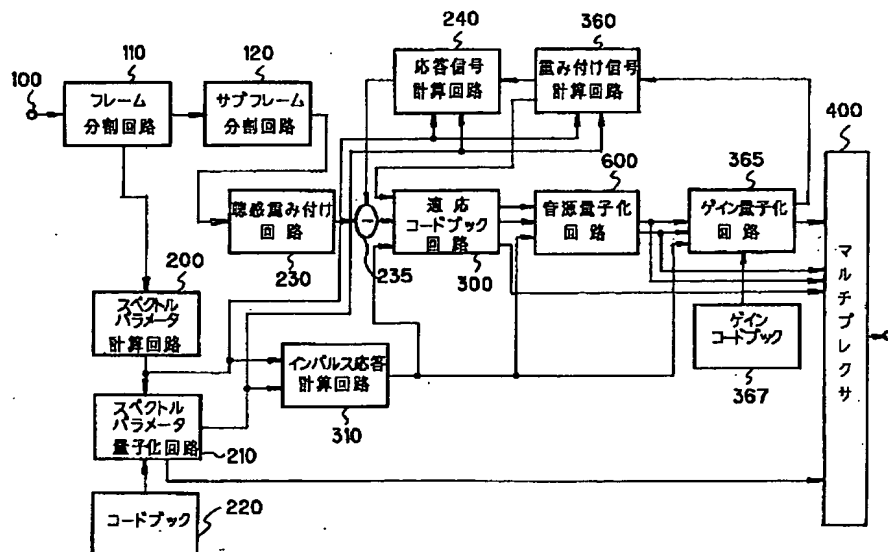




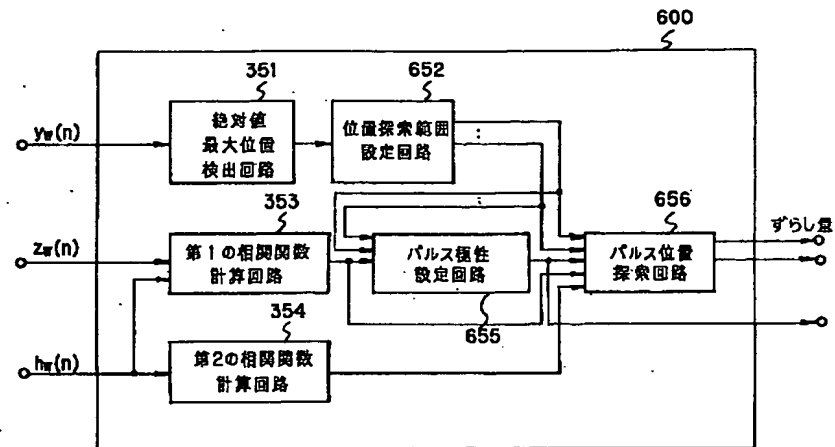
【図8】



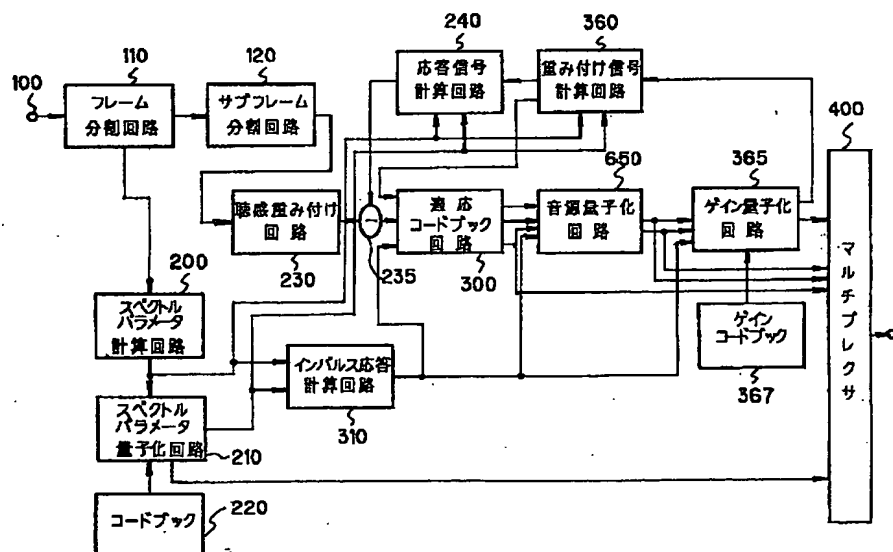
【図9】



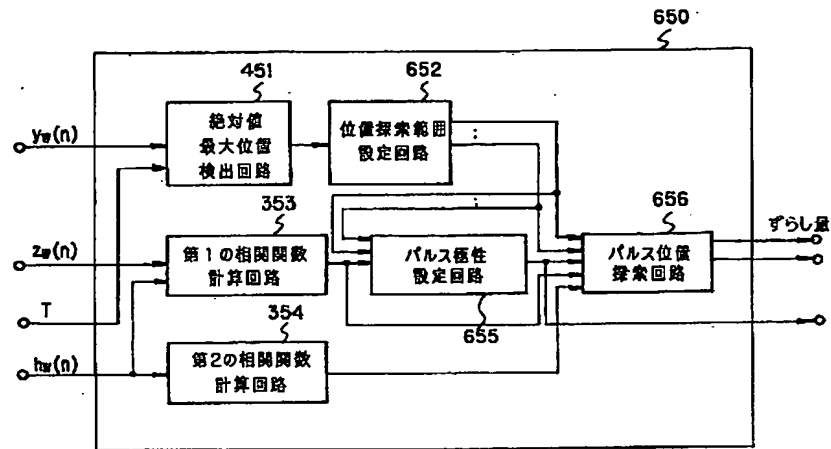
【図10】



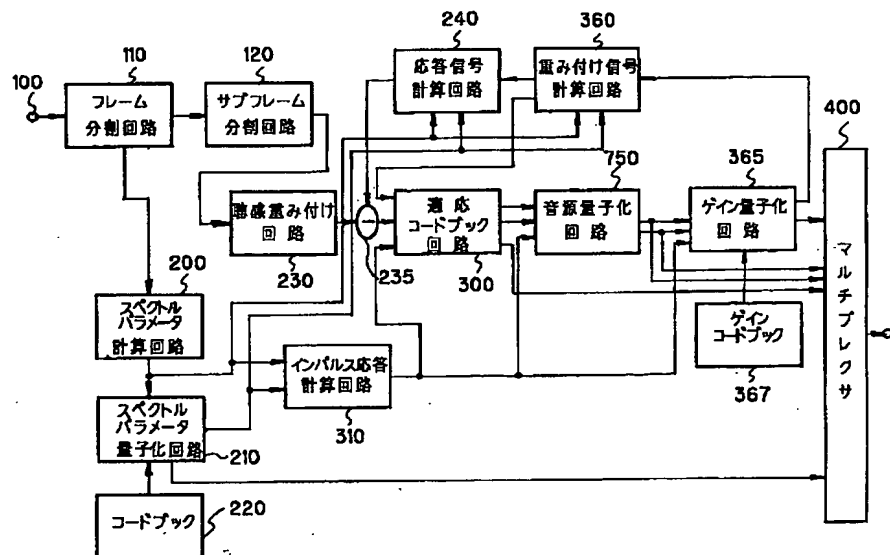
【図11】



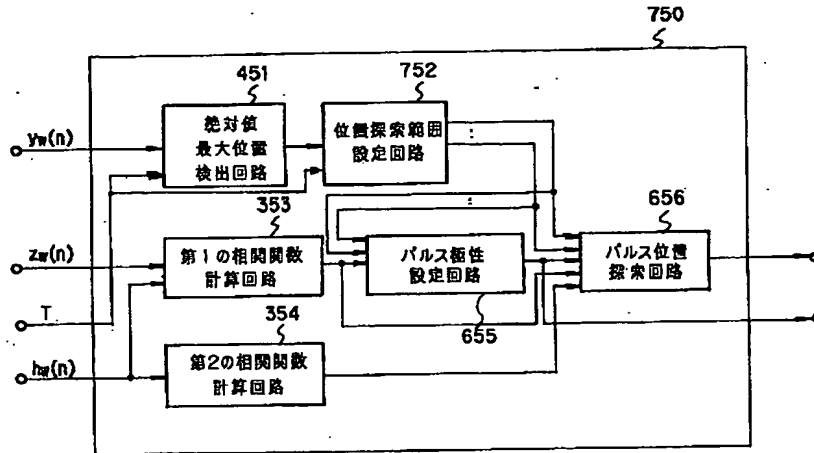
【図12】



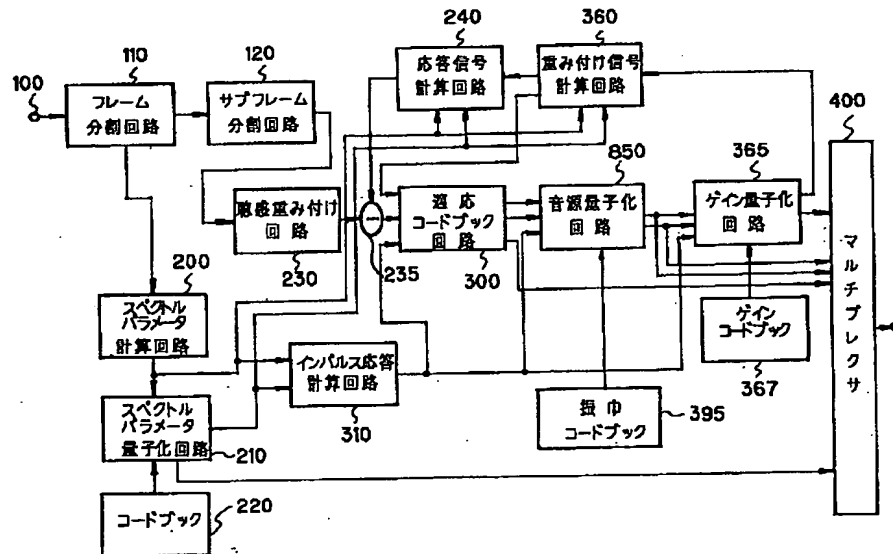
【図13】



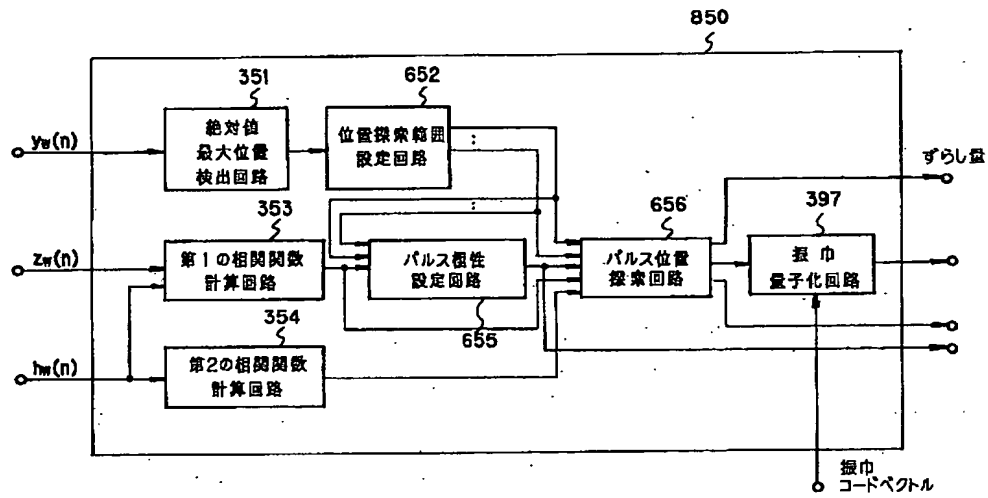
【図14】



【図15】



【図16】



【図17】

